Validity of performance indicators for assessing prescribing quality

Pont, LG; Denig, P; van der Molen, T; van der Veen, WJ; Haaijer-Ruskamp, FM

Published in:
European Journal of Clinical Pharmacology

DOI:
10.1007/s00228-003-0696-x

Publication date:
2004

Eur J Clin Pharmacol (2004) 59: 833– 840
DOI 10.1007/s00228-003-0696-x

PHARMACOEPIDEMIOLOGY AND PRESCRIPTION

**Lisa G. Pont · Petra Denig · Thys van der Molen
Willem Jan van der Veen · Flora M. Haaijer-Ruskamp**

Validity of performance indicators for assessing prescribing quality: the case of asthma

Received: 12 February 2003 / Accepted: 24 September 2003 / Published online: 18 November 2003

Abstract *Objectives*: The aim of this study was to assess the concurrent validity between the identification of sub-optimal treatment based on clinical information and computer generated indicators. Indicators that are associated with sub-optimal treatment in one of the four steps of asthma management were assessed.
Design: The ability of each indicator to identify patients with sub-optimal asthma treatment from computerised general practitioner (GP) prescription records was assessed by comparing them with the results of an individual patient assessment using clinical data.
Setting: Chronic asthma patients ($n = 146$) registered with 16 Dutch GPs.
Main measures: The sensitivity and positive predictive value (PPV) of each performance indicator was determined.
Results: The step-1 indicator, focusing on patients not prescribed a short-acting β-agonist, had an acceptable sensitivity (0.86), but a low PPV (0.52). The two step-2 indicators, targeting under-prescription of inhaled corticosteroids, had sensitivities of 0.74 and 0.37 and PPVs of 0.46 and 0.71, respectively. The step-3 indicator, which targeted under-dosing of inhaled corticosteroids, had a sensitivity of 0.07 and a PPV of 0.2. The fourth indicator, focusing on under-prescription of long-acting β-agonists, could not be validated due to inadequate numbers of patients with severe asthma in our study sample.
Discussion: None of the indicators investigated was considered valid for assessing prescriber performance, despite having good face and content validity. Performance indicators that have not been validated can only provide a broad-brush approach for assessing prescribing quality and should be used with extreme caution.

Keywords Quality indicators · Drug therapy · Asthma

L. G. Pont
Department of Clinical Pharmacology,
University of Groningen,
The Netherlands

P. Denig
Department of Clinical Pharmacology,
University of Groningen,
The Netherlands

T. van der Molen
Department of General Practice,
University of Groningen,
The Netherlands

W. Jan van der Veen
Registration Network Groningen,
Department of General Practice,
University of Groningen,
The Netherlands

F. M. Haaijer-Ruskamp (✉)
Department of Clinical Pharmacology,
University of Groningen,
Ant. Deusinglaan 1, 9713 AV Groningen,
The Netherlands
E-mail: F.M.Haaijer-Ruskamp@med.rug.nl
Tel.: +31-50-3632810
Fax: +31-50-3632812

Introduction

Quality assessment and improvement is receiving attention world wide. Information on health-care quality is being demanded by policy makers, health-care professionals and the general public alike [1, 2, 3, 4, 5, 6, 7]. With the majority of doctor–patient encounters in general practice resulting in a prescription for drug treatment, the quality of prescribing in general practice is an important issue [8]. Prescribing indicators for general practice have been used in several countries, for example, in the US, UK, Australia and Netherlands [7]. They are likely to keep a central role in programs to optimise care and are used for a range of purposes. The most important are (1) identification of patients receiving sub-optimal care, (2) monitoring of change or assessing the outcome of interventions and (3) identification of poor performers for purposes of postgraduate education or accountability and regulation [7, 9]. However, as illustrated by Pringle et al., creating

meaningful indicators from accurate data is a challenge [7]. This is especially the case for conditions that require a complex treatment regime according to severity, such as asthma.

Asthma is a chronic condition affecting more than 8% of adults in Western Europe. It is commonly treated in general practice [10]. Since the early 1990s, international asthma treatment guidelines have been available, which provide clear, well-accepted recommendations regarding optimal asthma management. Asthma is a variable disease, where the severity is determined by symptoms and lung function. The international consensus distinguishes four different levels of severity, each demanding a different treatment with medicines, i.e. optimal asthma pharmacotherapy is a step-wise process with treatment differing for each of the four asthma severity classes [11, 12, 13].

Various indicators have been developed for assessing the quality of asthma treatment in general practice, based on the recommendations in the international guidelines [14, 15, 16, 17, 18, 19, 20, 21]. Most indicators are derived from easily accessible prescription data. Commonly used asthma indicators focus on the proportion of patients prescribed a particular medication, such as inhaled corticosteroids or short-acting β-agonists, without taking into account differences in asthma severity [16, 17, 22]. An international collaboration developed a set of five asthma-prescribing indicators, attempting to differentiate among asthma severity classes using drug-based proxies. These indicators were used to evaluate an international intervention for improving asthma care in general practice [18, 23, 24]. The face and content validity of these indicators was considered good [23, 24]. Face validity indicates that the indicator appears to measure what it purports to and content validity means that the indicator is considered to be relevant. However, later work showed little correlation between indicators designed to target the same sub-optimal treatment pattern [19], suggesting that face and content validity alone may not be adequate for determining indicator validity. Concurrent validity, that is, comparison with a reference measure, is needed to fully determine the validity of performance indicators. The aim of this study was to assess the concurrent validity between the identification of sub-optimal treatment based on clinical information and computer generated indicators.

Materials and methods

Study population and recruitment procedure

The Registration Network Groningen (RNG) is a general practice database in the northern Netherlands. At the time of the study, the RNG included 30,486 patients registered with 16 general practitioners (GPs). Participating GPs use the database in clinical practice in place of paper medical records. Due to this, all prescriptions can be linked to the indication asthma, as diagnosed by the GP.

All patients aged 18–49 years with an asthma medication (Anatomical Therapeutic Chemical classification–ATC [25] group R03) or an asthma contact (International Classification of Primary Care-ICPC [26] code R96) during 1997 were selected from this database. Patients who received asthma medications for non-asthma indications were excluded, as well as patients who were no longer registered with an RNG doctor. Data from 1997 were used for recruitment to ensure that patients had chronic asthma in the study period. Eligible patients were invited to an individual study appointment by their GP, who unlocked the "key" that is used in the database to safeguard anonymity. A reminder letter was sent to non-respondents within 3 months of the initial invitation. Anonymous data for eligible non-responders were used for comparison regarding age, sex and asthma medication use. The local medical ethics committee approved the study, and informed consent was obtained from each participant.

Validation process

We validated indicators that targeted sub-optimal treatment patterns related to four steps included in the 1997 National Institutes of Health (NIH guidelines) [12].

Step 1. Under-treatment of short-acting inhaled β-agonists for all asthma patients
Step 2. Under-treatment of inhaled corticosteroids for mild, moderate or severe persistent asthma
Step 3. Inadequate dose of inhaled corticosteroids for moderate or severe persistent asthma
Step 4. Under-treatment of inhaled long-acting β-agonists for severe persistent asthma

Every patient was classified as being sub-optimally treated within every severity level of asthma in two ways: (1) using the computer-based prescribing indicators (the "indicators") and (2) using information on currently used medication in relation to symptoms and lung function assessed in a study appointment (the "individual patient assessment"). The latter classification was used as the reference value for the validation, i.e. the validation process checked to what extent the identification of sub-optimal treatment based on clinical information can be reproduced by computer generated indicators.

Indicators

The indicators validated in this study have been previously developed and used internationally for assessing the quality of prescribing for chronic adult asthma [18, 20, 24]. Detailed information on the development and use of the indicators has been published elsewhere [19]. In this study, indicators were calculated using data from the RNG database from 1999–2000. For each participating asthma patient, prescription data for the 12-month period prior to the study appointment were obtained retrospectively from the GP database. The indicators, based on the sub-optimal treatment patterns related to four steps included in the 1997 National Institutes of Health (NIH guidelines) [12] were calculated as follows.

Step 1 indicator. The patient is not prescribed an inhaled short-acting β-agonist during the 1-year period prior to the study appointment
Step 2A indicator. The patient is not prescribed an inhaled corticosteroid during a 1-year period prior to the study appointment
Step 2B indicator. The patient is not prescribed an anti-inflammatory agent (either corticosteroid or cromoglycate), but is prescribed an amount of inhaled short-acting β-agonist for daily use over a 1-year period prior to the study appointment

Step 3 indicator. The patient is prescribed a low-dose inhaled corticosteroid (≤ 400 μg budesonide or equivalent daily, Table 1) in combination with an inhaled short-acting β-agonist more than once daily over a 1-year period prior to the study appointment

Step 4 indicator. The patient is prescribed a high-dose inhaled corticosteroid (>600 μg budesonide or equivalent, Table 1), and an inhaled short-acting β-agonist at least twice daily over a 1-year period prior to the study appointment, but no long-acting β-agonist

The step-2B, step-3 and step-4 indicators incorporate the average daily amount of short-acting β-agonist over a 1-year period per patient. This is used in these types of indicators as a proxy for asthma severity, assuming that higher β-agonist use is associated with more severe asthma.

Individual patient assessment

The information needed for the reference measure on sub-optimal treatment patterns was collected in an individual study appointment. Each participating patient attended a single study appointment (May 2000–December 2000) with a trained research assistant. During this appointment, a questionnaire on symptoms and medication was completed. The questions referred to symptoms of asthma, such as dyspnoea and night symptoms, during the previous week and to drug use in the previous week, asking for the name of the drug and the number of doses used per day for every drug. Patients were asked to bring all current asthma medications to the study appointment and were asked about having spare bronchodilator inhalers, which may 'distort' prescription patterns, as seen in computerised databases. Each patient's forced expiratory volume at 1 s (FEV_1) was determined according to the standards of the American Thoracic Society, using a Microlab 3300 spirometer (Micro Medical Ltd., Rochester, Kent UK). For each participant, the best of three readings was used. Each patient's treatment regime was then classified as being either optimal or sub-optimal, taking into account severity and current medication, according to the explicit criteria specified in the NIH guideline (Table 1). For example, a patient with symptoms more than twice a week, but not daily, or exacerbations that may affect activity or night time symptoms more than twice a month, but not weekly, or a sub-optimal, but ≥80% predicted FEV_1, was classified as a class-2 patient. According to the NIH instructions, the presence of at least one of the features of severity is sufficient to place a patient in that severity class. An individual was assigned the most severe class in which a feature occurred.

Sample size

We defined that a valid indicator would require both a sensitivity and positive predictive value (PPV) of at least 0.7 for use in quality assessment. With a 95% confidence interval of 0.1, we calculated that at least 84 asthma patients would be needed to validate each indicator [27]. The RNG database included 369 patients meeting our inclusion criteria, and we decided to invite all eligible patients to participate in the study to ensure that our sample size requirements were met.

Analysis

Since each indicator can be viewed as a type of diagnostic test that identifies sub-optimal treatment, we assessed indicator validity in terms of sensitivity and PPV. These were calculated according to Altman [28]. The sensitivity of an indicator denotes the proportion of cases of sub-optimal treatment found in the individual assessment method, which is also identified by the indicator. The PPV is the proportion of cases identified by the indicator that is confirmed in the individual assessment.

Three of the indicators (step 2B, step 3 and step 4) incorporated levels of daily short-acting β-agonist use as a proxy for asthma severity. The association between short-acting β-agonist use and asthma severity was explored using rank correlation (Kendall's tau). A perfect correlation yields a Kendall's tau of 1, while a Kendall's tau of 0 indicates no relationship. In order to assess if the cut-off levels used in each indicator could distinguish among the

Table 1 Severity classification criteria and pharmacotherapy recommendations from the 1997 National Institutes of Health asthma guideline. *FEV_1* forced expiratory volume at one second, *PEV* peak expiratory flow

Asthma severity class	Criteria for classifying asthma severity*	Main pharmacotherapy recommendations
Class 1. Mild intermittent	Symptoms: two times a week or less. Asymptomatic and normal PEF between exacerbations. Exacerbations brief. Nighttime symptoms: two times a month or less. Lung function: FEV_1 or PEF ≥ 80% predicted	Short-acting inhaled β-agonist as needed.
Class 2. Mild persistent	Symptoms: > two times a week but < once a day. Exacerbations may affect activity. Nighttime symptoms: > two times a month. Lung function: FEV_1 or PEF ≥ 80% predicted	Short-acting inhaled β-agonist as needed and low† dose inhaled corticosteroid
Class 3. Moderate persistent	Symptoms: daily. Daily inhaled short-acting β-agonist use. Exacerbations affect activity. Exacerbations two times a week or more. Nighttime symptoms: > once a week. Lung function: FEV_1 or PEF > 60% and < 80% predicted	Short-acting inhaled β-agonist as needed and medium†† dose inhaled corticosteroid or low dose inhaled corticosteroid and inhaled long-acting β-agonist
Class 4. Severe persistent	Symptoms: continual. Limited physical activity. Frequent exacerbations. Nighttime symptoms: frequent. Lung function: FEV_1 or PEF 60% predicted or less	Short-acting inhaled β-agonist as needed and high††† dose nhaled corticosteroid and inhaled long-acting β-agonist and/or ipratropium bromide

*The presence of one of the features of severity is sufficient to place a patient in that category. An individual should be assigned to the most severe grade in which any feature occurs

†Low-dose inhaled corticosteroid is 200–400 μg budesonide (or equivalent) daily

††Medium-dose inhaled corticosteroid is 400–600 μg budesonide (or equivalent) daily

†††High-dose inhaled corticosteroid is >600 μg budesonide (or equivalent) daily

relevant asthma severity classes, we performed a receiver operating curve (ROC) analysis. On ROC curves, the optimal cut-off points can be found at the point where the curve begins to flatten.

Results

Of the 369 eligible patients identified from the database, 146 were willing to participate in the study. Patient characteristics and medication use are shown in Table 2. Current medication determined at the study appointment was higher than that obtained from the prescription database. Current short-acting β-agonist use was reported by 101 (69.2%), inhaled corticosteroids by 96 (65.8%), cromoglycates 8 (5.4%), long-acting β-agonists by 23 (15.8%) and ipratropium bromide by 19 (13.0%) participants. Of the patients not using a short-acting β-agonist, nine reported use of ipratropium bromide, and another five reported use of formoterol.

Non-respondent characteristics

Data from the GP database showed no difference with respect to gender between non-respondents and participating patients (58.2% and 57.8% female, respectively). Participating patients were slightly older (39.8 years versus 35.8 years, $P < 0.05$). There were no significant differences between participating and non-responding patients in the mean volume prescribed per patient for inhaled short-acting β-agonists, inhaled corticosteroids, inhaled anticholinergics and oral salbutamol.

Validation of the indicators

Table 3 shows the sensitivity and PPVs for each indicator. The validity of the step-4 indicator could not be determined due to an insufficient number of patients with persistent severe asthma (severity class 4).

The step-1 indicator had a sensitivity of 0.86, showing that prescription data correctly identified 86% of all patients without a short-acting β-agonist, as identified in the individual assessment. The PPV of 0.52 indicates that just over half of the patients identified by the indicator as not having a short-acting β-agonist prescribed during the study period actually did not have medication at the study appointment. The others did have this medication with them at the study appointment, either from earlier prescriptions or from prescriptions by other prescribers.

The step-2A indicator identified 74% of patients who were not prescribed an inhaled corticosteroid and who, according to their severity, should be prescribed this drug group. However, of all patients found by the indicator without an inhaled corticosteroid, less than half (PPV = 0.46) were confirmed by individual assessment. The step-2B indicator had a slightly lower sensitivity than the step-2A indicator, despite targeting the same sub-optimal treatment pattern. The majority (71%) of patients identified by this indicator as in need of an inhaled corticosteroid were confirmed upon the individual assessment. The remaining 29%, upon individual assessment, were either currently using an anti-inflammatory agent or were not high users of their short-acting β-agonist or, according to their severity classification, did not need an anti-inflammatory agent.

The step-3 indicator was the poorest performing indicator. This indicator was designed to detect patients with severity class 3 or higher who were prescribed a sub-optimal inhaled corticosteroid dose. However, it detected only 7% of the patients found upon individual assessment to have an inadequate corticosteroid dose. Furthermore, just 20% of the patients identified by the indicator with a sub-optimal dose were confirmed as having a sub-optimal corticosteroid dose upon individual assessment.

Short-acting β-agonist use as a proxy for asthma severity

While higher short-acting β-agonist use was significantly associated with increasing asthma severity (Kendall's tau = 0.307, $P < 0.001$), the wide standard deviations show considerable overlap in short-acting β-agonist use among the different asthma severity levels (Table 4).

Table 2 Characteristics of participating patients

General ($n = 146$)	Computerised general practitioner's database	Individual patient assessment
Mean age (years) ± SD	39.8 ± 8.3	39.8 ± 8.3
Number of females (%)	85 (58.2%)	85 (58.2%)
Mean lung function (% predicted ± SD)		83.5 ± 17.9
Medications	Number of patients (%)	
Inhaled short-acting β-agonists	86 (58.9%)	101 (69.2%)
Inhaled corticosteroids	90 (61.6%)	96 (65.8%)
Inhaled cromoglycates	6 (4.1%)	8 (5.4%)
Inhaled long-acting β-agonists	27 (18.5%)	23 (15.8%)
Inhaled ipratropium bromide	18 (12.3%)	19 (13.0%)
Asthma severity		Number of patients (%)
Class 1		51 (34.9%)
Class 2		9 (6.2%)
Class 3		70 (47.9%)
Class 4		16 (11.0%)

Table 3 Sensitivity and positive predictive values for each indicator (95% confidence intervals)

Indicator	Relevant asthma severity classes targeted by each indicator (number of patients)	Sensitivity	Positive predictive value
Step-1 indicator	All severity classes ($n=146$)	0.86 (0.71–0.95)	0.52 (0.38–0.65)
Step-2A indicator	Severity classes 2, 3 & 4 ($n=95$)	0.74 (0.57–0.88)	0.46 (0.33–0.60)
Step-2B indicator	Severity classes 2, 3 & 4 ($n=95$)	0.37 (0.19–0.58)	0.71 (0.42–0.92)
Step-3 indicator	Severity classes 3 & 4 ($n=86$)	0.07 (0.00–0.32)	0.20 (0.00–0.72)
Step-4 indicator	Severity class 4 ($n=16$)	Not validated due to inadequate sample size	

Table 4 Mean volume short-acting β-agonist amount prescribed per patient per day. Mean volume is presented in defined daily doses (DDDs) [27]. 1 DDD is 800 μg salbutamol (or equivalent); for two patients, the mean volume could not be calculated due to incomplete information

Asthma severity	Number of current users (individual patient assessment)	Mean volume (DDDs) per patient per day ± SD (computerised prescription data)
Class 1	23	0.24 ± 0.31
Class 2	9	0.46 ± 0.60
Class 3	51	0.60 ± 0.60
Class 4	16	1.14 ± 0.98

The step-2B indicator used a cut-off point for daily short-acting β-agonist use (200 μg of salbutamol daily or equivalent) to identify patients with a severity class of more than one from prescription data. As seen in Fig. 1a, this cut-off point correctly classified only 58.7% of all patients. The step-3 indicator used a cut-off point of high daily short-acting β-agonist use (400 μg salbutamol daily or equivalent) to identify patients of a severity class of more than two; however, this cut-off point correctly classified only 43.9% of patients (Fig. 1b). Neither curve flattens considerably at any point, indicating that no optimal short-acting β-agonist levels could be identified.

Patients were asked during the study appointment about having spare bronchodilator inhalers. More than half of the patients being prescribed a short-acting β-agonist reported using multiple inhalers simultaneously (55.7%). Of these, 50.7% reported having one spare inhaler, 34.3% kept two spares and the remaining 14.9% kept between three and five extra short-acting β-agonist inhalers.

Discussion

In this study, we evaluated the concurrent validity of five performance indicators derived from computerised prescription data for assessing the treatment of asthma. Two general indicators, which looked at the total number of patients prescribed short-acting β-agonist or inhaled corticosteroids, had reasonable sensitivities, but low PPVs. The two indicators that incorporated proxies for asthma severity had much lower sensitivities and varied PPVs.

While the response rate in this study was rather low, there is no reason to believe that this would affect the outcomes of our validation study. An adequate number of patients was recruited to validate four of the five indicators. No difference was seen between participants

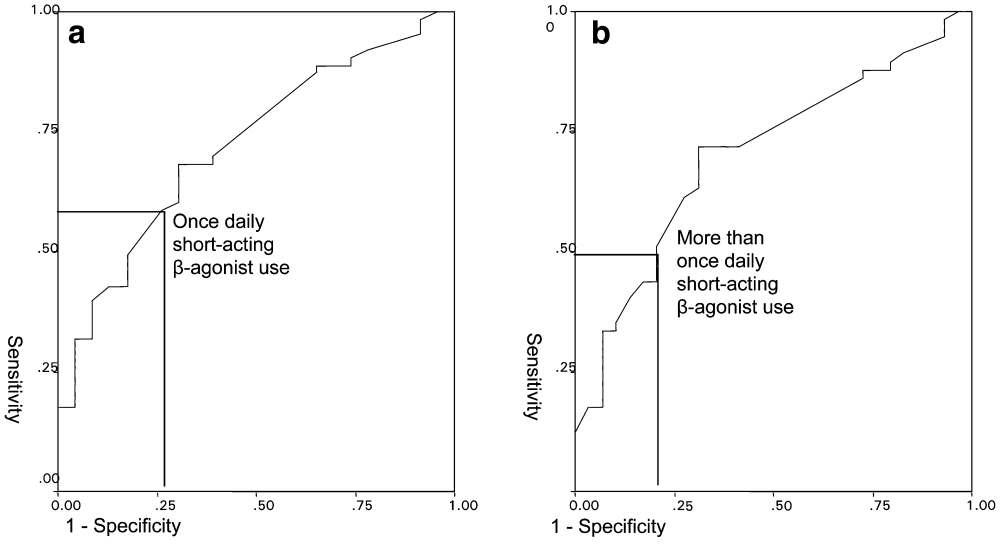

Fig. 1 Receiver operator curves (ROC) of mean defined daily doses of short-acting β-agonist use per day for identification of short-acting β-agonist cut-off points distinguishing among different asthma severity classes. **a** Identification of patients in severity classes greater than one. **b** Identification of patients in severity classes greater than two

and non-respondents with respect to gender and the use of the different asthma medications, although participating patients were slightly older.

We used a general practice database for calculating the prescribing indicators. This database gives an accurate account of all treatments prescribed by a GP to patients diagnosed with asthma. In the Netherlands, asthma medications are available only on prescription, and, in general practice, patients register with a single GP. Thus, the database contains complete records regarding the asthma management of GPs. Pharmacy records or claims data without information on disease are considered less reliable in identifying patients with asthma [29]. A limitation of using a GP database is that it does not register prescriptions from specialists. In the Netherlands, patients referred to a specialist may initially receive a prescription from the specialist, but generally return to their GP for further medication supply. For chronic medication, this is expected to result in a slight underestimation of medication when looking at a GP database during a 1-year period.

A GP prescription database can never fully reflect actual use of medications. It overestimates actual use when patients do get prescriptions, but are not collecting or using the medication. However, it may underestimate actual use when patients use their medication intermittently, and, therefore, do not need a new prescription during a 1-year period. In the Netherlands, one prescription may be valid for a maximum of a 3-month supply, but for medication that is used intermittently (''as needed''), a prescription may cover a longer period. This is expected to be the case for some relief medication, such as inhaled short-acting β-agonists, but it is also known that asthma patients sometimes use their inhaled corticosteroids intermittently. This could explain the finding that the number of patients identified from the GP database on certain medications during the study period was somewhat lower than as determined during the study appointment with the patients themselves. Self-reported medication use, by patient interview, has been shown to be a reliable method of obtaining current medications [30, 31, 32]. In this study, the accuracy of self-reported medication use was further improved by having patients bring their asthma medications to the study appointment. This would include specialist-prescribed medication as well as medication prescribed prior to the study period, but still used.

A limitation of this study is that the clinical assessment of treatment quality referred to the previous week, while prescribing indicators were based on prescribing in the previous year. This may have led to an underestimation of the validity of the prescribing indicators, since asthma severity might have changed during the preceding year. However, four of the indicators (step1, step 2A, step 2B and step 3) identify sub-optimal prescribing, in particular under-treatment, for several severity steps. Sub-optimal prescribing for step 1 includes, by definition, sub-optimal prescribing for steps 2, 3 and 4. This implies that changes in severity are covered by these indicators, with the exception of the step-4 indicator, which could not be validated due to lack of data.

The sensitivity and PPVs required for validity vary depending on the purpose for which the indicator is to be used. If an indicator is being used to identify patients with potential treatment problems, then it is important that the initial screening process identifies as many patients as possible who then undergo further, more accurate testing. Thus, for this purpose, indicators should have a high sensitivity. Of the five indicators validated in this study, both step 1 and step 2A indicators could be considered valid screening indicators; though still far from ideal, no better alternatives exist. The step-1 indicator identified 86% of all patients without an inhaled short-acting β-agonist, while the step 2A indicator detected 74% of patients not prescribed an inhaled corticosteroid who should have one according to their severity classification.

If an indicator is being used to monitor changes in drug utilisation or to assess the outcome of an intervention, then one may be satisfied with identifying only a selection of the patients of interest. In this case, the PPV becomes relatively more important than the sensitivity. Thus, if one is satisfied with monitoring changes or intervention outcomes in a sub-sample of patients, the step-2B indicator could be considered valid. Although this indicator identified less than half of all patients needing an inhaled corticosteroid, more than 70% of the patients identified were confirmed as needing an inhaled corticosteroid upon the individual assessment.

One of the common uses of prescribing indicators is to assess the performance of doctors. To be considered valid for assessing performance, an indicator needs to identify as many sub-optimally treated patients as possible while being certain that the all the patients identified are actually being treated sub-optimally. That is, both the sensitivity and PPVs should be sufficiently high. Clearly, none of the indicators based on prescription data validated in this study can be considered valid for this purpose.

A major problem affecting the validity of three of the indicators studied (step 2B, step 3 and step 4) was the inadequacy of using the prescription volume of short-acting β-agonist as proxy of asthma severity. While we observed a significant relationship between short-acting β-agonist use and asthma severity class, the range of use was extremely wide, and no clear cut-off points could be found that distinguished clearly among the different severity classes. More than half of the patients reported using multiple short-acting β-agonist inhalers simultaneously. Given the role of these medications as symptom relievers, it is understandable that patients may keep spare inhalers to ensure that one is always readily available. When calculating the average amount of short-acting β-agonist use over a 1-year period, patients using multiple inhalers simultaneously could appear to be high users if they have received a high number of repeat prescriptions during the study period, while, in reality, they may be moderate or even low users.

One may question whether the low sensitivity and PPV will also be found in other asthma populations. Of

course, all validation scores are determined by the underlying treatment pattern. Differences may be found, for example, when β-agonist use is more indicative of actual asthma severity than in the studied asthma population. However, this seems hardly likely, since treatment patterns in this population are much in line with that found in other countries in Europe [18].

This study emphasises the need for more information regarding the validity of the indicators calculated from computerised prescription data currently in use for assessing treatment quality. Computerised data may be easily available, however, in many instances, they may not provide enough information to judge a doctor's performance. This is, in particular, the case in complex diseases, such as asthma. Face and content validity, on which validation has centred to date [33, 34, 35], are not adequate substitutes for concurrent validity, checking if an indicator adequately describes what can be observed in actual clinical practice.

Prescribing indicators calculated from other databases than a GP database will probably have lower validity estimates than found in this study. If only prescribing data and no indication is available (as in sales data, claims data or pharmacist records), misclassification of patients occurs, undermining the possibility of a valid assessment of prescribing quality for a specific disease [29]. With the computerisation in general practice, however, the availability of GP databases and data on prescribing linked to indication increases fast in Europe.

Before an indicator is used, information on its sensitivity and PPV should be known. In the case of asthma and other complex diseases, prescribing indicators can only be used as a broad-brush approach to treatment quality when additional information on asthma severity class and actual use of asthma medications by the individual patient is lacking. This is even more of the case when no data are available that may link patients over time to prescribing behaviour or in the absence of a diagnosis. None of the commonly used asthma indicators calculated from prescription databases that were evaluated in this study can be considered a valid tool for assessing the quality of doctors' performance in the management of asthma.

With the growing international interest in performance indicators, this study emphasises the importance of indicator validation and underlines that performance indicators should only be used with caution.

Acknowledgements We wish to thank Annet Nicolai for her valuable assistance with the spirometry and patient assessment and Professor Dirkje Postma for her help in preparing this manuscript.

References

1. Spoeri RK, Ullman R (1997) Measuring and reporting managed care performance: lessons learned and new initiatives. Ann Intern Med 127:726–732
2. Brook RH, McGlynn EA, Shekelle PG (2000) Defining and measuring quality of care: a perspective from US researchers. Int J Qual Health Care 12:281–295
3. Sheldon T (1998) Promoting health care quality: what role performance indicators. Qual Health Care 7[Suppl]:S45–S50
4. Poloniecki J (1998) Half of all doctors are below average. BMJ 316:1734–1736
5. Davis RM (1999) Measuring the performance of public health agencies. Government, like doctors and hospitals, should meet quality standards. BMJ 318:889–890
6. Grol R, Wensing M (2001) Implementation: effective change in patient care. (In Dutch: Implementatie: effectieve verandering in de patientenzorg.) Elsevier, Maarssen
7. Pringle M, Wilson T, Grol R (2002) Measuring "goodness" in individuals and healthcare systems. BMJ 325:704–707
8. Dutch Association of the Research-based Pharmaceutical Industry (2002) Nefarma annual report 2000.
9. Nadzam DM (1991) Development of medication-use indicators by the Joint Commission on Accreditation of Healthcare Organizations. Am J Hosp Pharm 48:1925–1930
10. Rabe KF, Vermeire PA, Soriano JB, Maier WC (2000) Clinical management of asthma in 1999: the Asthma Insights and Reality in Europe (AIRE) study. Eur Respir J 16:802–807
11. National Institute of Health (1992) International consensus report on diagnosis and treatment of asthma. Clin Exp Allergy 22[Suppl1]:1–72
12. National Institutes of Health (1997) Expert panel report 2: Guidelines for the diagnosis and management of asthma. NIH publication number: 97–4051. National Institutes of Health, Bethesda
13. National Institutes of Health. (1995) Global strategy for asthma management and for asthma management and prevention NHLBI/WHO workshop. NIH publication number: 95–3659. NHLBI/WHO, Bethesda
14. Aveyard P (1997) Assessing the performance of general practices caring for patients with asthma. Br J Gen Pract 47:423–426
15. Shelley M, Croft P, Chapman S, Pantin C (2000) Is the quality of asthma prescribing, as measured by the general practice ratio of corticosteroid to bronchodilator, associated with asthma morbidity? J Clin Epidemiol 53:1217–1221
16. Hallas J, Hansen NC (1993) Individual utilization of anti-asthma medication by young adults: a prescription database analysis. J Intern Med 234:65–70
17. Legorreta AP, Christian Herman J, O'Connor RD, Hasan MM, Evans R, Leung KM (1998) Compliance with national asthma management guidelines and specialty care: a health maintenance organization experience. Arch Intern Med 158:457–464
18. Lagerlov P, Veninga CC, Muskova M, Hummers-Pradier E, Stalsby LC, Andrew M, Haaijer-Ruskamp FM (2000) Asthma management in five European countries: doctors' knowledge, attitudes and prescribing behaviour. Drug Education Project (DEP) group. Eur Respir J 15:25–29
19. Veninga CC, Denig P, Pont LG, Haaijer-Ruskamp FM (2001) Comparison of indicators assessing the quality of drug prescribing for asthma. Health Serv Res 36:143–161
20. Wahlstrom R, Lagerlov P, Lundborg CS, Veninga CCM, Hummers-Pradier E, Dahlgren LO, Denig P (2001) Variations in general practitioners' views of asthma management in four European countries. Soc Sci Med 53:507–518
21. Lagerlov P, Loeb M, Andrew M, Hjortdahl P (2000) Improving doctors' prescribing behaviour through reflection on guidelines and prescription feedback: a randomised controlled study. Qual Health Care 9:159–165
22. Gaist D, Hallas J, Hansen NC, Gram LF (1996) Are young adults with asthma treated sufficiently with inhaled steroids? A population-based study of prescription data from 1991 and 1994. Br J Clin Pharmacol 41:285–289
23. Veninga CC, Lagerlov P, Wahlstrom R, Muskova M, Denig P, Berkhof J, Kochen MM, Haaijer-Ruskamp FM (1999) Evaluating an educational intervention to improve the treatment

of asthma in four European countries. Drug Education Project Group. Am J Respir Crit Care Med 160:1254–1262
24. Veninga CC, Denig P, Zwaagstra R, Haaijer-Ruskamp FM (2000) Improving drug treatment in general practice. J Clin Epidemiol 53:762–772
25. WHO Collaborating Centre for Drug Statistics Methodology. (1999) ATC index with DDDs. World Health Organization, Olso
26. World organisation of family Doctors (WONCA) (2000) ICPC-2: international classification of primary care, 2nd edn. Oxford Medical Publications, Oxford
27. Bland M (2000) An introduction to medical statistics, 3rd edn. Oxford University Press, Oxford
28. Altman DG (1997) Practical statistics for medical research. Chapman & Hall, London
29. Pont LG, van der Werf GT, Denig P, Haaijer-Ruskamp FM (2002) Identifying general practice patients diagnosed with asthma and their exacerbation episodes from prescribing data Eur J Clin Pharm 57:819–825
30. de Jong-van den Berg LT, Waardenburg CM, Haaijer-Ruskamp FM, Dukes MN, Wesseling H (1993) Drug use in pregnancy: a comparative appraisal of data collecting methods. Eur J Clin Pharmacol 45:9–14
31. Van den Brandt PA, Petri H, Dorant E, Goldbohm RA, Van de Crommert S (1991) Comparison of questionnaire information and pharmacy data on drug use. Pharm Weekbl Sci 13:91–96
32. West SL, Savitz DA, Koch G, Strom BL, Guess HA, Hartzema A (1995) Recall accuracy for prescription medications: self-report compared with database information. Am J Epidemiol 142:1103–1112
33. Cantrill JA, Sibbald B, Buetow SA (1998) Indicators of the appropriateness of long term prescribing in general practice in the United Kingdom: consensus development, face and content validity, feasibility, and reliability. Qual Health Care 7:130–135
34. Campbell SM, Cantrill JA, Roberts D (2000) Prescribing indicators for UK general practice: Delphi consultation study. BMJ 321:425–428
35. Hanlon JT, Schmader KE, Samsa GP, Weinberger M, Uttech KM, Lewis IK, Cohen HJ, Feussner JR (1992) A method for assessing drug therapy appropriateness. J Clin Epidemiol 45:1045–1051